Spending Wisely?
How Resources Affect Knowledge Production in Universities

March 1, 2012

Abstract

Every year billions of dollars are spent on research grants to produce new knowledge in universities. However, as grants may also affect other research funding the effects of financial resources on knowledge production remains unclear. To uncover how financial resources affect knowledge production we study the effects of research spending itself. Utilizing the legal constraints on university spending from an endowment we develop an instrumental variables approach. Our approach instruments for university research spending with time-series variation in stock prices interacted with cross-sectional variation in initial endowment market values for research universities in the United States. Our analysis reveals that research spending has a substantial positive effect on the number of papers produced, but not their impact. We also demonstrate that research spending effects are quite similar at private and public universities.

Keywords: University, Government Grants, Research, Knowledge Production

JEL Classifications: H5, I2, O3

1 Introduction

The federal government spends billions of dollars each year on programs designed to produce new knowledge in universities. Public investments in sponsoring basic research are frequently argued to be central to the process of economic growth, and necessary for United States universities to retain international leadership in basic science.[1] Echoing these sentiments the National Academy of Sciences' *Gathering Storm* report called for a doubling of federally funded R & D in physical sciences over the next seven years (National Academy of Sciences, 2007). Moreover, as stimulus programs enacted to address the recent financial crisis have included a substantial increase in spending for basic science, public support for university research grant programs has increased dramatically.[2]

While the importance of federal funding for the financing of university research is clear; whether financial resources significantly increase knowledge production is subject to debate.[3] Adams and Griliches (1998) and Jacob and Lefgren (2011) find little evidence that research grant funding has positive effects on knowledge production. In contrast, Payne and Siow (2003), Adams (2009), and Gurmu, Black, and Stephan (2010) find positive effects of research grants on knowledge production in universities.

One important feature that divides recent studies is how the research grants they examine are allocated. For example, Payne and Siow (2003) examine the effects of research grant recepit from politically motivated earmarks and find significant positive effects on

[1]Seminal work by Lucas (1988) and Romer (1990) provides the conceptual basis for the role of knowledge production in economic growth. Highly influential empirical studies of the knowledge spillovers of universities include Jaffe (1989) and Zucker, Darby, and Brewer (1998). More recent work by Aghion, Boustan, Hoxby, and Vandenbussche (2009) and Kantor and Whalley (2009) also examine spillovers from universities, and Furman and Macgarvie (2007) examine the spillovers from basic science laboratories. Examples of recent work measuring the effects of university resources on student outcomes include Bound and Turner (2007), Bound, Lovenheim, and Turner (2009), and Bettinger and Long (2009a, 2009b)

[2]For example, the American Recovery and Reinvestment Act of 2009 (ARRA) allocates $3 Billion to the National Science Foundation, representing an increase of 50% over the NSF's annual budget of $6 Billion. Similarly, the ARRA allocates $10 Billion to the National Institute of Health, representing an increase of more than 30% of the NIH's annual budget of $30 Billion.

[3]Early work suggested that research grant funding has a positive effect on knowledge production (See Jacob and Lefgren (2011) for an excellent survey), however more recent work that carefully addresses the endogenous nature of government research funding presents mixed findings.

knowledge production. In contrast, Jacob and Lefgren (2011) examine the effects of competitively awarded NIH research grants and find little effect on knowledge production.

While differing returns to resources could explain the conflicting findings, there is another possibility. Researchers who receive politically motivated earmarks and researchers who receive competitively awarded research grants likely have very different alternative funding options. As Jacob and Lefgren (2011) show, researchers with promising projects who narrowly miss obtaining a highly competitive research grant are able to obtain funding from another source. In contrast, as politically motivated earmarks do not crowd out other university funding (Payne, 2001) their effects may be closer to those of financial resources alone. Thus, the contrasting findings could be due to differing fiscal impacts of grant receipt, even if the return to the financial resources spent is the same.

In this paper we attempt to reconcile the conflicting findings in prior work by estimating the effect of research spending on university knowledge production directly. Our approach addresses two important challenges. First, by estimating the return to research *spending*, rather than research grant income, our approach estimates the effect of financial resources alone. Second, research spending and knowledge production covary at the university level for a variety of reasons. Therefore, simple correlations are unlikely to reveal the effect of research spending on knowledge production alone. By exploiting potentially exogenous variation in research spending in universities, we attempt to isolate the elasticity of research spending on knowledge production. As we estimate the research spending effect, rather than the research grant effect, our estimates reveal the effects of financial resources themselves.

Our strategy is to exploit variation in university research spending due to the impact of stock market shocks on university endowment values. The legally-mandated and formulaic nature of spending from university endowments presents a particularly compelling instrument for university research spending. Because universities hold endowment resources in trust, they are legally bound to spend a fixed portion of the market value of the securities that they hold. Thus, exogenous stock market shocks will affect research spending across universities differentially depending on the size of the university endowment. Our instrument utilizes the cross-sectional variation in initial endowment market values across universities in the United States interacted with stock market returns to isolate variation in university income that is exogenous to unobserved research productivity in a particular university.

To conduct our analysis we use a previously under-explored data on university research spending and knowledge production covering the 96 leading research universities from 1981 to 1996. We use newly available data on the basic knowledge production of leading research universities compiled by Adams and Clemmons (2008) together with university spending, income, and endowment data from the Higher Education General Information Survey (HEGIS) and Integrated Post-Secondary Education Data System (IPEDS) collected by the US Department of Education.

We first examine the effect of research spending on the quantity of research produced. We find significant effects for basic knowledge. We find a research spending elasticity of about 1 on the number of papers produced. In addition, we find little evidence that research spending reduces applied knowledge production measured by patents, suggesting little trade-off between the production of basic and applied knowledge.

Next we examine how research spending affects the impact of the research conducted, measured by future citations to academic papers and patents. Marginal research spending may result in more – but lower impact – publications if researchers (and funders) are citation optimizing. We find some evidence of such a trade-off. For basic knowledge, our baseline instrumental variables estimates reveal a negative relationship between research spending and publication impact. While some of these estimates are not precise, the overall pattern of our results indicates a negative relationship between research spending and publication impact. Our results suggest that marginal increases in research spending result in the completion of marginal research projects that citation maximizing researchers would not have pursued. In sum, while our estimates are only indicative of a quality-quantity trade off in knowledge production, they clearly rule out positive effects of research spending on the impact of the knowledge produced.

Our final set of results examine whether research spending is more or less effective in privately or publicly controlled institutions. The unusually high degree of autonomy universities have from government control may be an important factor in the relative performance of US universities. Indeed, recent work by Aghion, Dewatripont, Hoxby, Mas-Colell, and Sapir (2009) (ADHMS) has presented evidence that the elasticity of local innovation with respect to research grant income is larger for more autonomous public universities. Our results do not indicate significantly larger returns to research spending at private universities, suggesting that governance structure plays little role in the direct effects of university research spending.

We also provide several robustness checks of our baseline estimates, particularly focusing on whether our causal estimates of the effect of research spending on knowledge production might be spurious. As our strategy exploits the differential effects of time-series variation in stock prices there is a potential concern that differential trends in unobservable determinants of knowledge production across universities could threaten our identification assumption. For example, recent research has argued that innovation at the frontier is becoming increasingly difficult (Jones, 2009) and information technology diffusion has had the greatest effect on research productivity at middle-ranking institutions (Agrawal and Goldfarb, 2008). One implication of these findings may be that secular increases in research productivity are likely to be smaller at leading research universities that also likely have the largest endowments. Comfortingly, we find little evidence that our central results can be explained by differential trends in unobservable determinants of knowledge production across universities.

The remainder of the paper proceeds as follows. In section 2, we outline our empirical approach to quantify the impact of university research spending on knowledge production. Section 3 describes our data and the descriptive statistics of key variables. In Section 4, we present our main results. Section 5 examines various robustness checks to probe the validity of our approach. Section 6 concludes.

2 Empirical Approach

Our empirical strategy is to estimate the effect of research spending on knowledge production using plausibly exogenous variation in spending due to endowment value shocks. In particular, we instrument for research spending in different universities with time-series variation in stock prices interacted with cross-sectional variation in the initial level of the endowment market value across universities. Our instrumental variables procedure accounts for both the fiscal impacts of endowment income and the endogenous nature of the allocation of research funding. We outline how our strategy addresses both issues in this section.

2.1 Structural Relationships

Consider the following structural equation of the effect of research spending on knowledge production,

$$Y_{it} = \beta R_{it} + \alpha_{yi} + \gamma_{yt} + \epsilon_{it}, \tag{1}$$

where Y_{it} is knowledge production in university i in year t, R_{it} is research spending in university i in year t, α_{yi} are time-invariant unobserved determinants of knowledge production in university i, γ_{yt} are time varying determinants of knowledge production at time t, and ϵ_{it} are other unobserved determinants of knowledge production in university i in year t. The parameter β is the effect of research spending on knowledge production. The simplest strategy to estimate (1) would be by ordinary least squares (OLS). However, OLS estimates are likely to be biased as research expenditure is likely to be positively correlated with unobserved determinants of knowledge production. Our empirical analysis begins with estimating a model similar (1) by OLS, only without the α_{yi} control, so that we can better understand what potential sources of bias might matter for the estimates we present.

To address the concern with bias in OLS estimates of β, exogenous variation in research spending is required. To obtain identification a commonly applied strategy is to utilize exogenous variation in research grant receipt. However, variation in research grant receipt alone will not identify the spending effect when research grants attract or displace other funding. To see this let G_{it} be exogenous variation in research grant income in university i in time t. The relationship between research grant income and research spending is given by,

$$R_{it} = \delta G_{it} + \alpha_{ri} + \gamma_{rt} + u_{1it}. \tag{2}$$

We can then specify the reduced form relationship between research grant income and knowledge production as,

$$Y_{it} = \beta\delta G_{it} + \alpha_{yi} + \gamma_{yt} + \beta\alpha_{ri} + \beta\gamma_{rt} + u_{2it}. \tag{3}$$

From model (3) we are able to estimate the reduced form effect of research grant income and obtain an estimate of $\beta\delta$. The reduced form estimate does not separately identify both the research spending effect (β), from the fiscal impact of the research grant (δ). The research grant effect $\beta\delta$ will exactly identify the spending effect only if research grant income G_{it} does not crowd out, or crowd-in, other resources (i.e. if $\delta = 1$). We estimate a two stage model that estimates both δ and β to identify the effects of financial resources on knowledge production alone.

2.2 Econometric Models

We implement three approaches to estimate the effect of research spending on knowledge production. First, to provide a baseline we estimate (1) by OLS without university characteristics or fixed effects. To address bias arising from time invariant differences in unobserved determinants of knowledge production (i.e. γ_{yt}) we then estimate a first-differenced version of (1). Finally, to address the endogenous nature of research spending we estimate an instrumental variables version of the first-differenced model.

Our first differences analysis estimates β with the following equation,

$$\Delta Y_{it} = \beta_1 \Delta R_{it} + \tau_t + v_{1it}, \tag{4}$$

where ΔY_{it} is the first difference in knowledge production in university i in period t, ΔR_{it} is the first difference in research spending in university i in period t, τ_t is a set of year fixed effects, and v_{1i} is the error term. We estimate the model in first-differences as many of the unobserved components of knowledge production that are likely correlated with research spending, such as the presence of highly productive faculty or advanced scientific laboratories, are time invariant. Estimating the models in first-differences means that the models are not identified off of this potentially problematic cross-sectional variation.[4]

[4]The downside to estimating the model in first-differences is that if much of the variation in ΔR_{it} within a university is driven by measurement error, our estimate of β would be attenuated towards zero.

While our first differences strategy addresses time-invariant sources of bias, it does not address the issue that changes in research spending within a university may be endogenously related to changes in unobserved university research productivity. For example, research grants are likely to be awarded to researchers with the most promising new projects. This would suggest that our estimate of β would be biased upwards even with a first differenced model, as v_{1it} would be positively correlated with ΔY_{it}. Conversely, if highly productive faculty face lower costs in financing their work they may be able to fund new projects with a low probability of success. This would suggest that our estimate of β would be biased downwards even with the first-differenced model, as v_{1it} would be negatively correlated with ΔY_{it}. Thus, in principal the bias could go in either direction.

Our main empirical strategy attempts to isolate potentially-exogenous sources of variation in research spending, ΔR_{it}. We instrument for changes in research spending by exploiting the differential impact of changes in stock prices across universities in which endowment revenue plays a more or less significant role in funding research spending. In particular, we instrument for ΔR_{it} in equation (4) with the following first-stage regression:

$$\Delta R_{it} = \delta_1 \Delta S_{t-1} * E_{i,1981} + \tau_t + v_{2it}, \tag{5}$$

where ΔR_{it} is the research spending in university i in period t, $\Delta S_{t-1} * E_{i,1981}$ is the first difference in stock prices in year $t-1$ (ΔS_{t-1}) interacted with the market value of the endowment in university i in 1981 ($E_{i,1981}$), τ_t is a set of year fixed effects, and v_{2i} is the error term. Our identifying assumption is that, absent stock price changes, knowledge production in universities with large and small endowments would have grown at similar rates.

Before continuing it is useful to be clear what the research spending effect captures. The research spending effect jointly captures the return to many inputs into the knowledge production process, from scientific equipment and laboratory space, to graduate student and faculty research time. While it would be highly relevant to know the return to each input separately we focus on return to total financial resources as this is the policy relevant parameter in our context. We also choose to specify our models at the aggregate university level, rather than per faculty member. We do this to allow research spending to affect

---

Fortunately, our instrumental variables strategy addresses both the endogeneity of, and measurement error in, research spending to achieve a consistent estimate of β.

the size of the research university sector, as well as the productivity of researchers in the sector. As such, our estimates capture both responses.

It is also useful to clarify the exact parameter we seek to estimate and how it differs from other broader effects of university research. Our approach focuses on estimating the direct effect of university research spending on knowledge production in universities. We do not seek to capture any of the spillover effects of university research that require development, adoption or investment responses by the private sector. As such our parameter is quite different from the full social return studies reviewed in Alston et al. (2000) that capture a broad range of spillover effects. This distinction matters because Alston et al. (2000) show the full social return to university research takes substantial time to manifest, up to 20 years in many cases.[5] As we seek to measure the direct effect of university research on university knowledge production alone we follow Payne and Siow (2003) and Jacob and Lefgren (2011) in studying the relatively short time horizon termed the "gestation period" by Alston el al. (2000).

2.3 Research Design: University Endowment Management and Spending Practices

The intuition behind our identification strategy is straightforward. Universities spend a fixed fraction of the market value of their endowments in any year because of legal constraints on the spending of endowment resources held in trust. As Ehrenberg (2000 and 2009) notes, universities follow a rule of spending 4 to 5 percent of the market value of their endowments each year.[6] Since universities generally follow their own stable payout

[5]Whether significant spillover effects exist with an academic community remains subject to debate. See for example, Azoulay, Graff Zivin, and Wang (2010), Waldinger (2012), and Borjas and Doran (2012).

[6]The fixed-spending rule emerged in the early 1970s as a result of efforts to maximize the long-term value of endowment portfolios and to increase their long-term effectiveness as a source of revenue (Yoder's 2004). This policy comes from an influential 1969 Ford Foundation report that concluded that universities could indeed spend capital gains by using a total return spending policy. The report also recommended that universities follow a total-return spending policy based on a three-year moving average of their endowments' market values, regardless of whether endowment income came from capital gains or distributions. Yoder (2004, 10) notes that differences across institutions in their target spending rates are small, differences in the rate of return they experience may well be larger. Indeed, universities with higher

rule and all have different endowment values, then exogenous stock market shocks will lead to variation in the amount of endowment income each university will spend in any one year. As stock market shocks and the level of the initial endowment are exogenous to trends in knowledge production across universities, this variation provides a compelling source to identify the effects of overall university expenditures on knowledge production.

Endowment income can be spent on a range of inputs that affect university knowledge production. New equipment could be purchased, more graduate students enrolled or funded, faculty teaching loads could be reduced with endowment income. For example, Kantor and Whalley (2011) show that universities enrol more graduate, but not undergraduate, students in response to positive endowment shocks. While university endowment shocks provide a compelling source of exogenous variation in university expenditure the types of expenditures funded by endowments may not be identical to those funded by a NIH research grant for example. However, as the prior literature remains mixed on whether university research spending has any effect on knowledge production we view identifying the causal effect of endowment driven spending to be policy relevant.

2.4 Threats to Identification

Our identifying assumption is that, absent stock price changes, knowledge production in universities with large and small endowments would have grown at similar rates. This is reasonable since both national stock prices and the initial market value of a university's endowment are not affected by, and should not be correlated with, changes in a university's unobserved research productivity. Of course, universities with large and small endowments may differ in other ways that are likely to affect scientific productivity. Any such differences that are time-invariant will be differenced out, and not contribute to identification in our first differences approach. Only differential trends in scientific productivity across these universities would be a threat to the validity of our instrumental

SAT admission score experienced a 1.4 percent greater return on their endowments from 1992 to 2005, primarily due to differences in portfolio allocation (Lerner, Schoar and Wang, 2008). Increases in portfolio allocation to alternative asset classes (i.e. hedge funds, private equity, etc.) largely occurred after our sample period. Lerner, Schoar and Wang (2008) note that in 1992 these types of assets accounted for only 1.1 percent of all assets, but grew to 8.1 percent in 2005.

variables strategy. We provide a variety of evidence in favour of our identifying assumption by estimating models which allow for other effects of stock prices. However, it is useful to consider cases where our identification assumption may be threatened.

First, it is possible that stock market shocks reflect economic shocks that affect universities differentially. For example, it could be the case that time-series variation in stock prices reflects time-series variation in productivity growth, perhaps from advances in information technology. As advances in information technology could affect high or low endowment universities differentially, we may estimate an effect where none was present. For example, as Agrawal and Goldfarb (2008) demonstrate, the effects of IT diffusion are particularly large for middle-tier universities. Similarly, recent trends in the role of teams in innovation may affect large universities, with likely large endowments, differentially (e.g. Wuchty, Jones, and Uzzi, 2007). To address these and related concerns we estimate models where we include a variety of linear trends at the state and university level, as well as falsify with future values of research spending.

Second, it is possible that stock market shocks affect universities differentially for reasons that have little to do with how much they spend on research. For example, stock market shocks may affect university based innovation through their effects on the financial resources of private sector collaborators.[7] If highly endowed universities are more likely to collaborate with private sector firms, this may undermine our identification strategy.[8] To address this and related concerns we estimate models where we allow changes in knowledge production in each university to be differentially correlated with changes in the stock market depending on the characteristics of the university.

In sum, while we cannot completely rule out the possibility that some of the effect reported below reflects time varying university-specific changes in unobserved scientific productivity, it appears that many sources of spurious correlation are accounted for.

[7]For example, private-public sector collaboration has been shown to be important in the innovation process in the case of drug discovery (Cockburn and Henderson, 1998).

[8]It is also possible that higher-quality universities hold a different portfolio of assets in their endowments (see Lerner, Schoar, and Wang, 2008). As higher-quality institutions are more likely to hold assets that are less correlated with stock market shocks, this may weaken our first stage for this group of universities.

2.5 Other Estimation Considerations

Clarity about the timing of our variables is especially important given the fact that we are identifying our parameter of interest off of changes in the variables over time. Many universities use the previous year's market value of endowment to determine how much is spent from the endowment in the next year. To be consistent with this fact we estimate the first stage of our IV models using one lag of stock market changes interacted with the initial endowment. In addition, the university knowledge production variable we use is reported based on calendar year activity, while university expenditure is reported on a fiscal year basis. To allow for university expenditure to have time to impact knowledge production we lag university spending by three survey years.[9] Thus, to take account of differences across the variables in the timing of reporting and behaviour we implement our first differences model in (4) as,

$$\Delta Y_{it} = \beta_1 \Delta R_{it-3} + \tau_t + \epsilon_{it}. \tag{6}$$

The first stage for the IV model above becomes,

$$\Delta R_{it-3} = \delta_1(\Delta S_{t-4} * E_{i,1981}) + \tau_t + \epsilon_{2it}, \tag{7}$$

where ΔR_{it-3} is the first difference in university research expenditure in university i lagged by three years, $\Delta S_{t-4} * E_{i,1981}$ is the first difference in the Standard and Poor's 500 stock index lagged four years (ΔS_{t-4}) interacted with the initial endowment level in university i ($E_{i,1981}$), τ_t is a set of year fixed effects, and ϵ_{2it} is the error term.

We do not include the main effect of stock market shocks in the model as the year fixed effects flexibly control for the time-series variation in the outcomes, capturing the main effect of stock market shocks on university knowledge production. Our central parameter of interest is β_1 which measures the effect of university research spending on knowledge production. To account for serial correlation in the outcomes and endowment income within a university we cluster the standard errors by university. We also note that as our model is just identified a meaningful test of over-identifying restrictions is not possible in our context.

[9] In an unreported analysis we have also examined the sensitivity of our results to estimating the models with a five year rather than three year lag structure. The results of this analysis are very similar to those with the three year lag we report and are available from the authors on request.

3 Data

To implement the analysis we require data on university expenditure and revenue, knowledge production, and an exogenous source of endowment variation. In this section we outline the data sources we use to conduct our empirical analysis.[10]

University Finances We obtain fifteen years of annual data on university expenditure, revenue, faculty, student, facilities, and ownership status from the Higher Education General Information Survey (HEGIS) and Integrated Post-Secondary Education Data System (IPEDS) for 1981 to 1996. The HEGIS/IPEDS data are a census of all four year colleges in the U.S. and reports information on revenue, expenditure, enrolment, and institutional characteristics from each university. We use HEGIS data until it was replaced with the IPEDS survey in 1984. We end our analysis in 1996 because the Department of Education has not released the college financial data for the 1997-2000 years, and the knowledge production data we use ends in 1999. Data on the market value of the university endowment is also collected, which is critical for our study. The cross-sectional distribution of initial endowment market values is displayed visually in Figure 1.

The primary variables obtained from HEGIS/IPEDS are research expenditures and endowment market values. Our research expenditure variable is based on the sum of (1) expenditures on sponsored research projects and (2) expenditures on research and teaching. As one of the component variables contains expenditures on teaching, in addition to research, we use the average time allocation of faculty between research and teaching to obtain our measure of research spending alone. We thus weight this total by the percentage of time the average faculty member spends on research activities within research and doctoral conferring institutions according to the National Study of Post-secondary Faculty in 1993 (NSPF93).[11]

[10]For further details on the construction of each variable and on the sample construction see the Data Appendix.

[11]The fraction of time allocated to research is 0.612. We use this national level adjustment to capture the average level of research spending in all universities in our sample. Of course, our estimates which is based on with-in university variation, does not depend on how we adjust the research expenditure and teaching measure to capture research alone.

University Knowledge Production We match the HEGIS/IPEDS data to the scientific output of universities from the NBER-Rensselaer Polytechnic Institute Scientific Papers Database (NBER-RES) (Adams and Clemons, 2008). This dataset contains information on annual counts of academic publications, all forward citations to academic publications published in a given year, as well as, collaboration ties, for all authors at the 110 leading research universities for papers published from 1981 to 1999. Our measure of citations per publication is the total future citations to academic publications published in a year divided by the number of publications in that year. We aggregate the NBER-RES data to the university-year level because our research spending from the IPEDS/HEGIS data measure is at the university level. The fact that the data are first available in 1981 determines the initial year in our analysis.

We also obtain data on patents and forward patent citations from the NBER Patent Database (NBER-PAT) (Hall, Jaffe, and Trajtenberg, 2001). These data are measured at the individual patent level, giving exact application and grant dates along with field and institutional information. We match these data to our sample of universities using a crosswalk developed by the United States Patent and Trademark Office to match universities to patent assignees. We collapse the data to patent and patent forward citation counts at the university-application year level. Our measure of citations per patent is simply the ratio of the number of forward citations to a patents granted in that year divided by the number of patents granted in that year.

Other University and Regional Micro Data Together the data from the HEGIS/IPEDS, NBER-RES, and NBER-PAT form the panel of universities that we use for our central analysis. We match this core data set to additional data from two further sources. First, we obtain data on private sector sponsored research income from the NSF Survey of Federal S & E Support to Universities, Colleges, and Non-profit Institutions for 1981 to 1996. Second, we obtain data on state population and per capita income from the 1980 Census.

Stock Prices We construct our instrument by interacting the initial university endowment market value in 1981 with the Standard & Poor's 500 Index in each year for each university. We normalize the S & P 500 Index so that the 1981 value is one. This normalization implies that changes in the index reflect changes in the market value of the 1981 endowment over time. As university expenditure are reported for the academic year from July to June, we use the average value of the Standard and Poor's Index over the same time period in order to align the timing of stock market shocks with university expenditures. The first difference in stock prices is stationary and displays little persistence, as we expect based on the well known random-walk property of stock prices. The time series of the level of stock prices is displayed visually in Figure 2.

Sample Construction The initial sample consists of the top 110 research institutions (i.e. universities, research institutes, and hospitals) as defined in the NBER-Rensselaer Polytechnic Institute Scientific Papers Database (NBER-RES). From this initial sample we drop research institutions for two reasons; either (1) they are not a research university and so are not reported in the HEGIS/IPEDS data; or (2) they are a research university, but there are data constraints for key variables.

We begin by dropping any institutions that are medical or are narrowly focused research institutes, such as oceanographic institutes. This results in six institutions being dropped from the sample. Specifically, we drop the following institutions: University of Texas Houston Health Science Center, Woods Hole Oceanographic Institute, University of Texas San Antonio Health Science Center, University of Texas Southwestern Medical Center of Dallas, Oregon Health Sciences University, Baylor College of Medicine. Next, when the remaining sample of 104 universities is matched to our sources of data there are seven universities with missing values for the base year endowment market value, or are missing multiple years of research spending. Rutgers University is missing for a large number of years, and because no sensible imputation can be utilized, it is dropped. There is also a small set of universities are missing the endowment market values in the base year, which is necessary for the instrument generation, and are therefore dropped. These institutions are: University of Connecticut, University of Kansas - Main Campus, State University of New York at Buffalo, Baylor University, University of Utah, Virginia Commonwealth University and Rutgers University. We also drop the one university system that does not have individual campus financial variable reporting for the market value of the endowment, the University of Texas, as only one university would be in the sample for the outcomes, but the financial data would refer to all the nine campuses and six health centers in the University of Texas system. The eight research universities we drop due to data constraints are generally smaller, less prominent, and publicly controlled. We are left with 96 out of the original top 110 non-profit research institutions for our analysis sample.

Due to the lag structure of the model we use data from 1985-1996 with 96 observations each year for a full sample of 1152 observations.

Descriptive Statistics Table 1 presents descriptive statistics in the base year (1981). Columns (1) and (2) show the means and standard deviations computed over all university observations dividing universities by baseline above or below median research spending levels. The comparison yields a number of interesting results. First, the quantities of publications and patents do differ significantly between above- and below-median research expenditure universities. However, the impact of the knowledge produced, measured by future citations per paper and citations per patent is very similar across universities with different levels of research spending. This cursory look at cross-sectional patterns would suggest a significant impact of university research spending on the number of papers and patents, but not their impact. Second, universities with above-median levels of spending receive significantly more research grant income. Third, there are also significant differences in university size, quality, faculty salary, and private sector collaboration between universities with above- and below-median research spending, but there is little difference in terms of the fraction public, or local economic characteristics. As university size, quality, and private sector collaboration are likely to affect university knowledge production independently of research spending, and are likely correlated with important unobservables, this comparison demonstrates the value of using an IV strategy to achieve a causal estimate of the impact of university research spending.

3.1 Endowment Income and Research Spending

In Column (1) of Table 2 we present the results from estimating the first-stage model in equation (7). The estimates in column (1) of Table 2 show that the coefficient on the interaction between initial endowment and stock market fluctuations one year prior results in a strong first stage. As noted above, universities often follow a 3 year smoothing rule to translate endowment market value into actual disbursements. In column (2) we dig more deeply into universities' spending policies to examine whether multiple lags of the endowment market value independently explain university research spending. Again, we find that the interaction between the first and second lag of stock returns and the initial endowment are highly statistically significant. In either case, the F-statistic on the excluded instruments in the first stage is well above the threshold level of 10 that has been established as key to reducing the finite sample bias inherent in IV methods (Bound, Jaeger, and Baker, 1995). We choose the single lag model for our baseline specifications as the first stage is stronger and we are able to estimate our models using more years of data.

3.2 Main Results

In Table 3 we present the central results of the paper. The results of a single regression are displayed in each column. In columns (1) to (3) we report results where the dependent variable is the level and first difference in number of academic publications at a university in a year. In columns (4) to (6) we report results where the dependent variable is the level and first difference in the average number of future citations to the academic papers published at a university in a given year per paper published.

Publications We first consider level OLS models of the relationship between university research spending and the quantity of knowledge production without any university controls in column (1) of Table 3. In column (1) we see that research expenditure is strongly correlated with academic publication production. However, as there are substantial concerns that universities with unobservable higher research productivity are able to attract higher levels of funding, in column (2) of Table 3 we present models where we estimate the model in first differences. The results in column (2) also reveal a correlation. Interestingly, the magnitude of the coefficient estimate in column (2) is far smaller than the estimate in column (1), which may indicate that universities with higher levels of time invariant research productivity are better able to attract resources. Alternatively, within university changes in research spending may be subject to substantial measurement error, attenuating the first differences estimates towards zero.

In column (3) of Table 3 we present our instrumental variables estimates. Our estimates indicate that research expenditure has a positive and statistically significant effect on the quantity of basic knowledge produced in a university. The effects appear to be both economically and statistically significant. The magnitude of our estimates indicates that a 1% increase in baseline research expenditure ($1.08 Million) increases the number of papers published by about 9 papers or 0.96% of baseline research output, for a research spending elasticity of about 1. Perhaps surprisingly, our instrumental variables estimates in column (3) are very similar to those in column (1), suggesting that the endogeneity of research spending with respect to the quantity of research is less of a concern than we might have suspected. However, they are quite different from those in column (2) though the difference-in-Sargan statistic does not indicate the FD-IV and FD-OLS estimates are statistically different.

How do our estimates in column (3) compare to prior work? Our point estimates are very close to Payne and Siow (2003) who find that an increase in earmarks for research funding of $1 Million increases the number of papers by about 10. In contrast, the implications of our IV estimates are quite different from the IV estimates of the effect of a NIH grant receipt on future publications reported in Jacob and Lefgren (2011) as they find no statistically significant effect of research grant receipt on the quantity of papers published. Thus, our findings indicate that the mixed findings in prior work could be due to NIH grants, but not politically motivated earmarks, crowding out other research funding.

Citations Per Publication Research projects on the margin of receiving funding may result in less significant discoveries, resulting in a trade off between the number and impact of publications. For example, citation maximizing researchers who experience an increase in research funding may add lower impact projects to their portfolio reducing the average impact of their completed projects. To explore this possibility we next examine the effect of research spending on one measure of the significance of the research discovery: future academic citations per paper.[12]

We first present level OLS regression models without any university controls in column (4) of Table 3. We see that the point estimate on research spending is positive, but statistically insignificant. However, when we estimate the models in first differences in column (5) of Table 3 we see that the sign of the relationship between research spending

[12]As our forward citation measure is computed at a point in time more recent publications will have less citations. Fortunately, our year fixed effects will control for differences in citation frequency by publication cohort.

and citations changes. In fact, the point estimate is now negative, indicating that when time invariant differences across universities are accounted for, research spending is negatively related to impact of the knowledge produced. In the last column of the Table we present our IV estimates of the effect of research spending on citations per publication. Again, our point estimates indicate that marginal changes in research spending reduce the impact of the knowledge produced by universities and the estimate is statistically significant at the 5% level. However, while the negative point estimates in column (6) are statistically significant, they are quite modest in magnitude. The estimates in column (6) suggest an elasticity of about -0.3. Again, the difference-in-Sargan test statistic indicates that the FD-OLS and FD-IV estimates are not statistically different.[13]

These findings echo some recent studies. Payne and Siow (2003) also find a negative relationship between research funding an citations per paper, however their IV estimates are statistically insignificant and smaller than the estimates reported in Table 3. Jacob and Lefgren (2011) present mixed and statistically insignificant IV estimates for the effect of NIH grant receipt on publication citations.

Patents and Patent Citations The results thus far have demonstrated that university research spending has a significant positive effect on the quantity of basic knowledge produced. We next examine the effect of university research spending on applied knowledge production.[14] The effect is priori ambiguous. It may be the case that researchers substitute away from applied research projects towards basic research projects reducing the amount of applied research completed. Alternatively, if basic and applied research are complements, then the additional production of basic research would lead to additional output of applied research. As many universities play a significant role in producing applied knowledge these effects are well worth examining.

In Table 4 we examine whether research spending affects applied research in a similar fashion to basic research. To do so we estimate the effect of university research spending

[13]In an (unreported) analysis we estimate similar models to equation (6) but examine a five year lag of research spending instead of the three year lag above. The pattern of results is very similar to those in Table 3, with larger point estimates in absolute value estimated with less precision.

[14]Cockburn, Henderson and Stern (1999) show that incentive structure of pharmaceutical research is consistent with complementarity between basic and applied knowledge production.

on the number of patents produced by a university, and the average number of forward citations to those patents. In the first two columns of Table 4 we see that the point estimates of the effect are all positive and statistically significant at the 5% level in columns (1) and (2).[15] The first difference OLS and IV point estimates in columns (3) and (4), while imprecisely estimated, are quite similar to those in columns (1) and (2). Again, the difference-in-Sargan test statistic indicates that the FD-OLS and FD-IV estimates are not statistically different. While the results are not conclusive, they do not indicate that a large basic-applied knowledge substitution effect is at work.[16]

In terms of the effect of research spending on the impact of the applied knowledge produced we obtain a similar pattern of point estimates to those above for basic science. Again, the sign of the point estimates depend on how the relationship is estimated. In columns (5) and (6) of Table 4 we see that the average number of forward citations to patents are positively related to research spending in the levels specifications, thought the point estimate is very small and not statistically significant. The first differences results in column (7) also display little relationship between the average number of forward citations to patents and research spending. In contrast, the IV point estimates in column (8) indicate a negative impact of research spending on the average number of forward citations to patents, though again the point estimate is not statistically significant. Interestingly, the difference-in-Sargan test statistic reveals that the FD-OLS and FD-IV estimates are not statistically different. In sum, the effects of research spending on applied knowledge production are broadly similar to those found for basic knowledge.

Crowding Out One important difference between our analysis and that of previous work is the focus on research spending effects rather than research grant effects. As the

[15]One potential issue with our patent outcomes that a university may issue zero patents in a given year. We address this issue by also reporting Tobit specifications in column (2) and (6) of Table 4. In general the Tobit and OLS estimates are quite similar. The fact that zero observations for patents comprise only 15.4 percent of the sample could account for this similarity.

[16]One potential concern with our measure of applied knowledge production is that for some university systems patents are assigned to the entire system and not allocated to the individual campuses. We allocate these system wide patents to individual campuses of the universities in the system based university with-in system shares of knowledge production. See the data appendix for details. In an (unreported analysis) we have also estimated these models with only universities that do allocate patents to campuses directly and obtained generally similar results to those reported in Table 4.

differences between these effects hinge upon whether a dollar of research income results in a dollar of research spending (δ in equation (2)), it is natural to ask whether endowment income crowds out (or crowds in) other sources of funding. In this subsection we test for a crowding response to endowment income fluctuations.

As we do not have data on endowment income distributed for research, we pursue a simple reduced form approach to testing for a crowding response.[17] As such, our estimates are more informative about the existence and direction of a meaningful crowding effect, rather than the magnitude of the effect. The reduced form model we estimate is simply to substitute $\Delta S_{t-4}{*}E_{i,1981}$ for R_{it-3} in equation (6) above with various external funding variables as the outcome measures.

We present evidence for the crowding effects of endowment income in Table 5. Each column presents the results for a different external source of research and development funding as the outcome variables. The two panels reflect differences in the lag length in the models between when the funding outcome and research income are measured. Each cell in the table reports the results for one regression where the funding outcome appears in the column heading and the lag length appears in the panel heading. The results reveal little evidence of a crowding effect of endowment income on other sources of funding at either time horizon.

While the lack of a crowding response here indicates that the spending and endowment income effects are likely to be similar, this is unlikely to be true in other contexts. For example, recent estimates indicate that federal funding leads to a \$0.33 increase in non-federal funding at U.S. universities (Blume-Kohout, Kumar, and Sood, 2009). A crowding out effect of this size implies that the research spending and research grant effects of knowledge production would differ by 33%. Thus, the differences between the research spending and grant effects are likely to be larger for many federal research grant programs than for endowment income.

[17] Unfortunately, data on endowment distributions by university function is not reported in the HEGIS/IPEDS data. Additionally, while a variable titled 'endowment income' is collected it does not include endowment distributions from the 'quasi-endowments' (i.e. capital gains) that are likely very important for our sample of leading research universities so we do not use it here.

3.3 Does The Form of Governance Matter?

One key difference in the structure of research universities between the United States and other advanced economies is the role of the private sector. Even within the United States the most prominent universities are often privately controlled. As universities with higher levels of autonomy may be able to allocate resources more productively, it is natural to wonder whether the particularly strong performance of US research universities is due primarily to private sector governance. Indeed, one policy under consideration in multiple countries involves providing universities more autonomy in an attempt to emulate the performance of leading private US universities. Recent work by Aghion, Dewatripont, Hoxby, Mas-Colell, and Sapir (2009) (ADHMS) finds important differences in the effectiveness of university research and development income on local innovation by the degree of university autonomy.[18]

In this section we examine whether the effects of research spending on knowledge differ by university control type. Our analysis differs from ADHMS (2009) in a number of ways. First, by examining differences in the effect of financial resources on knowledge production between publicly and privately controlled universities we consider a potentially broader range of autonomy than the within public sector differences examined by ADHMS (2009). Second, we examine the effect of research spending directly, rather than research grant income, as university governance structure might also affect financing constraints. Third, we examine the effect of university research spending on the knowledge produced by universities themselves, rather than the total knowledge produced in the local area.

Before discussing our results it is worth pointing out that it is not obvious which governance structure will yield greater returns to marginal increases in research spending. Privately controlled institutions may have higher returns to marginal spending if autonomy may allow them to fund particularly promising research projects. However, as Glaeser (2002) has pointed out, because private controlled institutions are largely faculty controlled and faculty value research over instruction, these universities may conduct more

[18]Consistent with a meaningful role for governance, Payne and Roberts (2010) find that research activity at public research universities does respond to performance measures. Similarly, Azoulay, Manso and Graff Zivin (2011) show that the incentives faced by grant recipients affect the quantity and direction of knowledge production.

research than a public university.[19] With diminishing returns to research, the return to a marginal increase in research spending may well be lower in private universities than in public universities. Therefore, the relationship between university governance structure and the effectiveness of marginal changes in research spending is an open question.

To examine whether there are differences in the effectiveness of university research spending on knowledge production we stratify the sample into public and private institutions. We then estimate our first differences model equation (6) by instrumental variables. We report the results of this exercise in Table 6. Again each main entry in each cell presents the results of a single regression. We first report the IV estimates with the first difference in publications as the outcome variable in columns (1) and (2) and the estimates with the first difference in publication citations in columns (3) and (4).

Table 6 reveals a number of interesting patterns. First, in comparing column (1) and (2) we can see that the point estimates do show some difference in the strength of the relationship between research spending and the quantity of knowledge produced by university control type. However, as the results for the publicly controlled institutions are relatively imprecisely estimated these differences are not statistically significant. We find very similar point estimates by university control type for publication impact in columns (3) and (4). Again, as the estimates are relatively imprecise, we do not regard these as meaningful differences. While the lack of significant differences may be driven by the imprecision in the estimates for public universities, the results do not point to substantively larger effects of research spending at private universities.

4 Robustness

In this section, we provide several robustness checks of our baseline estimates, particularly focusing on whether our causal estimates of the effect of research spending on knowledge production might be spurious. In the interest of brevity, we focus our discussion on the robustness of our main dependent variables, academic publications and citations to

[19] In our sample privately controlled universities spend about 13% more on research than publicly controlled institutions.

academic publications.

4.1 Alternative Instrumental Variables Specifications

In Table 7 we examine whether our results are robust to alternative specifications of the instrument. Figure 1 shows that the endowment market value distribution is highly skewed and one may be concerned that using the level of endowment market values might give disproportionate weight to the universities with the highest endowments. To address this issue we examine whether our results are sensitive to specifying the instrument as the log of the market value of the initial endowment rather than the level. We present the results in the second column of Table 7. While our point estimates are somewhat closer to zero than in the baseline, the central implications of the results above remain.

Moreover, as universities have substantial fixed costs, the effect of endowment market values on research spending, and thus on knowledge production, may be non-linear, with large and very large endowments leading to similar effects on innovation when stock prices rise. Motivated by these considerations, we report results with an alternative measure of endowment market values, where the endowment market values are top coded at the 95th percentile of the endowment distribution (the instrument is then constructed by interacting this measure with stock price) in column (3) of Table 7. The results, though slightly smaller in magnitude, are very similar to the baseline and remain statistically significant. Lastly, we examine whether our results are robust to alternative choices of the set of interactions in the first stage. As noted above many universities use a three year moving average of stock returns to determine their spending policy. We present estimates where we use the three year moving average first stage in column (4) of Table 7. Again the results are very similar to the baseline and remain statistically significant.

4.2 Exclusion Restriction

The exclusion restriction underlying our IV strategy is that absent stock price changes, universities with different levels of endowments would have experienced the same changes in knowledge production. In Table 8 we explore a variety of alternative specifications designed to investigate the validity of this identifying assumption.

We examine the evidence for two potential violations of our identification assumption. First, trends in unobservable determinants of university innovation may differ depending on the size of the initial endowment of the university. Universities with larger endowments are likely to be higher quality for example, and thus may be subject to different secular trends in knowledge production. To address this potential concern we estimate our IV models with alternative specifications that include controls for differences in underlying trends across universities. We begin by allowing for different linear trends in university knowledge production across states. We present the results of this analysis in Table 8 column (2). The results indicate that differing trends in knowledge production across universities in different states do not explain the central results. The point estimates are quite similar in magnitude to the baseline results, presented in column (1) of Table 8. Next we allow for university specific linear trends by including university fixed effects in the baseline first differenced model. We present the results of this model in column (3) of Table 8. The inclusion of university specific trends does little to alter the point estimate of the research spending elasticity. In fact, the point estimate is now larger than in the baseline model, and remains statistically significant at the 10% level despite the loss of precision. Given that universities differ on a number of unobservables measures that likely affect knowledge production trends, we regard this as an important specification check.

A final check we consider for our results being driven by differences in underlying trends across universities is to include a lead effect of our instrument. This is essentially a falsification analysis as there should not be a causal effect of *future* endowment income on knowledge production before it is received and used. The results in column (4) of Table 8 are comforting. There is little evidence of an effect of research spending before such an effect should occur. Furthermore, the main effect of research spending on knowledge production is very similar to the baseline estimate and remains statistically significant. While the point estimates of the effect of research spending on citations per publication do switch sign, the points estimate remains statistically insignificant. In sum, there is little evidence that our central results are driven by differential trends in unobserveable determinants of knowledge production across universities.

Second, we examine whether our results are robust to allowing knowledge production in different universities to be differentially correlated with stock market shocks. One potential concern with our identification strategy is that the direct effect of stock market shocks on knowledge production may differ across universities even if research spending has no effect on knowledge production. If, for example, other unobserved research spending by private sector collaborators is affected by the stock market and universities with large endowments are more likely to collaborate with the private sector, then our IV strategy would be weakened. To test for these possibilities we estimate various versions of the models in equations (6) and (7) where we allow the effect of stock market shocks to affect knowledge production depending on other time-invariant characteristics of universities.[20]

In columns (5) and (6) of Table 8 we allow for knowledge production in universities with above median levels of publications per faculty member and faculty size in the baseline year to be differentially correlated with stock market shocks. In either case, the coefficient remains statistically significant and similar in magnitude to the baseline specification. We then examine whether the differential direct effects of stock market shocks on knowledge production in universities that charge higher tuition (measured by baseline average tuition) or frequently collaborate with private sectors firms (measured by baseline number of publications with private sector collaborators) explain the results. Again, the point estimates are quite similar to our baseline estimate in column (1) and remain statistically significant. Thus, there is little evidence that differential exposure to stock market shocks across universities explains our central findings.

[20] Specifically, we extend models (6) and (7) as,

$$\Delta Y_{it} = \beta_1 \Delta R_{it-3} + \beta_2 (C_i * \tau_t) + \tau_t + \epsilon_{it}. \tag{8}$$

The first stage of the IV model then becomes,

$$\Delta R_{it-3} = \delta_1 (\Delta S_{t-4} * E_{i,1981}) + \delta_2 (C_i * \tau_t) + \tau_t + \epsilon_{2it}, \tag{9}$$

where ΔR_{it-3} is the first difference in university research expenditure in university i lagged by three years, $\Delta S_{t-4} * E_{i,1981}$ is the first difference in Standard and Poor's 500 stock index lagged four years (ΔS_{t-4}) interacted with the initial endowment level in university i ($E_{i,1981}$), ($C_i * \tau_t$) is the additional initial characteristic in university i (C_i) interacted with the year fixed effects (τ_t), τ_t is a set of year fixed effects, and ϵ_{2it} is the error term.

5 Conclusion

In this paper we have quantified the effect of research spending on knowledge production in universities in the United States. Our analysis reveals three main findings. First, we find a research spending elasticity of about 1 on the number of papers produced. Second, we find little evidence that research spending has a positive effect on the impact of the knowledge produced. In fact, the majority of our estimates show a negative relationship between marginal research spending and citations per paper. Third, we find little evidence that the effect of research spending on knowledge production is greater in private versus public universities at the margin. The results suggest that reforms to increase university autonomy at public universities to emulate private universities may have little effect on the effectiveness of university research spending.

While our analysis presents clear evidence on the effectiveness of university research spending, as with any empirical analysis, it has several limitations that suggest directions for future work. First, as returns to marginal research spending may well be different in universities in other countries with different institutional structures, future work examining the returns to research spending these contexts would be of interest. Second, as we lack data on the exact inputs used in the production of knowledge, future work investigating which inputs have the largest effects would be relevant for policy makers. More broadly, as crowding-out (crowding-in) effects are likely important in the delivery of public goods by many non-profit providers, our results demonstrate the value measuring spending effects directly.

6 References

Adams, James D. (2009) "Is the U.S. Losing Its Preeminence in Higher Education?," National Bureau of Economic Research, *Working Paper*, 15233.

Adams, James D. and J. Roger Clemmons (2008) "The NBER-Rensselaer Scientific Papers Database: Form, Nature, and Function," National Bureau of Economic Research, *Working Paper*, 14575.

Adams, James D. and Zvi Griliches (1998) "Research Productivity in a System of Universities," *Annales d'Economie et de Statistique*, ADRES issue, pp. 49-50.

Aghion, Philippe, Dewatripont, Mathias, Hoxby, Caroline M., Mas-Colell, Andreu, and Andre Sapir (2009) "The Governance and Performance of Research Universities: Evidence from Europe and the U.S.," National Bureau of Economic Research, *Working Paper*, 14851.

Aghion, Philippe, Boustan, Leah, Hoxby, Caroline M. and Jerome Vandenbussche (2009) "The Causal Impact of Education on Economic Growth: Evidence from U.S.," Harvard University, *Working Paper*.

Agrawal, Ajay and Avi Goldfarb (2008) "Restructuring Research: Communication Costs and the Democratization of University Innovation," *American Economic Review*, vol. 98(4), pp. 1578-90.

Alston, Julian M., Wyatt, T. J., Pardey, Philip G., Marra, Michele C., and Connie Chan-Kang (2000) "A Meta-Analysis of Rates of Return to Agricultural R& D: Ex Pede Herculem?" Research Report No. 113, Washington D.C.: International Food Policy Research Institute, 2000.

Azoulay, Pierre, Graff Zivin, Joshua S., and Jialan Wang (2010) "Superstar Extinction,?" *Quarterly Journal of Economics*, 125(2), pp. 549-589.

Azoulay, Pierre, Manso, Gustavo and Joshua Graff Zivin (2011) "Incentives and Creativity: Evidence from the Academic Life Sciences," *RAND Journal of Economics*, 42(3), pp. 527-554.

Azoulay, Pierre, Ding, Waverly, and Toby Stuart (2009) "The Effect of Academic Patenting on the Rate, Quality, and Direction of (Public) Research Output," *Journal of Industrial Economics*, 57(4), pp. 637-676.

Bettinger, Eric and Bridget Terry Long (2009a) "Addressing the Needs of Underprepared Students in Higher Education: Does College Remediation Work?," *Journal of Human Resources*, 44(3), pp. 736-771.

Bettinger, Eric and Bridget Terry Long (2009a) "Does Cheaper Mean Better? The Impact of Using Adjunct Instructors on Student Outcomes,"*Review of Economics and Statistics*, forthcoming.

Borjas, George and Kirk Doran (2012) "The Collapse of the Soviet Union and the Productivity of American Mathematicians," *Quarterly Journal of Economics*, forthcoming.

Bound, John and Sarah E. Turner (2007) "Cohort Crowding: How Resources Affect Collegiate Attainment," *Journal of Public Economics*, 91(5-6), pp. 877-899.

Bound, John, Lovenheim, Michael, and Sarah E. Turner (2009) "Why Have College Completion Rates Declined? An Analysis of Changing Student Preparation and Collegiate Resources," National Bureau of Economic Research, *Working Paper*, 15566.

Blume-Kohout, Margaret E., Kumar, Krishna B., and Neeraj Sood (2009) "Federal Life Sciences Funding and University R & D," National Bureau of Economics Research, *Working Paper* 15146.

Cockburn, Iain M., and Rebecca M. Henderson (1998) "Absorptive Capacity, Coauthoring Behavior, and the Organization of Research in Drug Discovery," *Journal of Industrial Economics*, 46 (2), pp. 157-182.

Cockburn, Iain M., Henderson, Rebecca M. and Scott Stern (1999) "Balancing Incentives: The Tension Between Basic and Applied Research," National Bureau of Economic Research, *Working Paper*, 6882.

David, Paul A., Hall, Bronwyn H., and Andrew A. Toole (2000) "Is Public R & D a Complement or Substitute for Private R & D? A Review of the Econometric Evidence," *Research Policy*, 29,pp. 497-529.

Ehrenberg, Ronald G. (2000) *Tuition Rising: Why College Costs So Much*, Cambridge, MA: Harvard University Press.

Ehrenberg, Ronald G. (2009) "Demystifying Endowments," Cornell Higher Education Research Institute, *Working Paper*, 118.

Furman, Jeffrey L., and Megan J. MacGarvie (2007) "Academic Science and the Birth of Industrial Research Laboratories in the U.S. Pharmaceutical Industry," *Journal of Economic Behavior & Organization*, 63(4), pp. 756-776.

Glaeser, Edward L. (2002) "The Governance of Not-For-Profit Firms," National Bureau of Economic Research, *Working Paper*, 8921.

Gurmu, Shiferaw, Black, Grant C. and Paula E. Stephan (2010) "The Knowledge Production Function for University Patenting," *Economic Inquiry*, 48(1), pp. 192-213.

Hall, Bronwyn, Adam B. Jaffe, and Manuel Tratjenberg (2001) "The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools," National Bureau of Economic Research, *Working Paper*, 8498.

Jacob, Brian and Lars Lefgren (2011) "The Impact of Research Grant Funding on Scientific Productivity," *Journal of Public Economics*, 95(9-10), pp. 1168-1177.

Jaffe, Adam B. (1989) "Real Effects of Academic Research," *American Economic Review*, 79(5), pp. 957-70.

Jones, Benjamin F. (2009) "The Burden of Knowledge and the Death of the Renaissance Man: Is Innovation Getting Harder?," *Review of Economic Studies*, 76(1), pp. 283-317.

Kantor, Shawn and Alexander Whalley (2011) "Do Universities Generate Agglomeration Spillovers? Evidence from Endowment Value Shocks," National Bureau of Economic Research, *Working Paper*, 15299, revised.

Lerner, Josh, Schoar, Antoinette, and Jialan Wang (2008) "Secrets of the Academy: The Drivers of University Endowment Success," *Journal of Economic Perspectives* 22 (Summer), pp. 207-222.

Lucas, Robert E. (1988) "On the Mechanics of Economic Development," *Journal of Monetary Economics*, 22(1), pp. 3-42.

Payne, A. Abigail (2001) "Measuring the Effect of Federal Research Funding on Private Donations at Research Universities: is Federal Research Funding More than a Substitute for Private Donations?" *International Tax and Public Finance*, 8(5-6), pp. 731-751.

Payne, A. Abigail and Aloysius Siow (2003) "Does Federal Research Funding Increase University Research Output?" *Advances in Economic Analysis & Policy*, 3(1).

Payne, A. Abigail and Joanne Roberts (2010) "Government Oversight of Organizations Engaged in Multiple Activities: Do Centralized Performance Schemes Encourage Quantity or Quality?," *Review of Economics and Statistics*, 92(1), pp. 207-212.

Romer, Paul M. (1990) "Endogenous Technological Change," *Journal of Political Economy*, 98(5), pp. S71-S102.

Salem, David (1992) "Endowment Management," in Deirdre McDonald Greene, *College and University Business Administration, fifth edition*, Washington, DC: National Association of College and University Business Officers.

Veugelers, Reinhilde and Frederick Van Der Ploeg (2008) "Reforming European Universities: Scope for an Evidence-Based Process," CESIFO, *Working Paper*, 2298.

Waldinger, Fabian (2012) "Peer Effects in Science: Evidence from the Dismissal of Scientists in Nazi Germany,?" *Review of Economic Studies*, forthcoming.

Wuchty, Stefan, Jones, Benjamin F., and Brian Uzzi (2007) "The Increasing Dominance of Teams in Production of Knowledge," *Science*, 316(5827), pp. 1036-1039.

Yoder, Jay A. (2004) *Endowment Management: A Practical Guide*, Washington, DC: Association of Governing Boards of Universities and Colleges.

Zucker, Lynne G., Darby, Michael R., and Marilynn B. Brewer (1998) "Intellectual Human Capital and the Birth of U.S. Biotechnology Enterprises," *American Economic Review*,88(1), pp. 290-306.

TABLE 1: Descriptive Statistics, 1981

All monetary values in 1996$ at the University Level	Full Sample (1)	Above Median Research Expenditure (2)	Below Median Research Expenditure (3)	(2)-(3) t-stat [p-value] (4)
<u>(1) Scientific Output:</u>				
Publications	1017	1461	572	8.35
	(684)	(665)	(319)	[0.00]
Citations Per Publication	10.7	10.8	10.6	0.05
	(4.5)	(4.3)	(4.8)	[0.82]
Patents	4.8	6.4	1.8	2.16
	(7.8)	(9.3)	(1.6)	[0.04]
Citations Per Patent	8.3	8.6	7.5	0.59
	(5.1)	(4.6)	(6.1)	[0.44]
<u>(2) University Expenditure ($1M):</u>				
Research Expenditure	108	161	57	11.88
	(67)	(57)	(20)	[0.00]
Total Expenditure	404	578	229	9.42
	(252)	(223)	(126)	[0.00]
Endowment Market Value in 1981	187	255	119	1.98
	(342)	(442)	(177)	[0.05]
<u>(3) University Research Funding</u>				
Federal Research Funding	106	155	52	5.51
	(103)	(120)	(33)	[0.00]
State Research Funding	12	17	5	3.28
	(18)	(22)	(10)	[0.00]
Private Industry Funding	6	9	3	3.69
	(7)	(9)	(4)	[0.00]
<u>(4) University Characteristics:</u>				
Number of Students	17170	22610	11489	6.23
	(10180)	(10075)	(6601)	[0.00]
Number of Faculty	813	1085	541	7.55
	(444)	(415)	(276)	[0.00]
Public	0.64	0.66	0.60	0.63
	(0.48)	(0.47)	(0.49)	[0.53]
U.S. News Quality Ranking	3.01	3.33	2.69	2.64
	(1.23)	(0.97)	(1.39)	[0.01]
Mean Faculty Salary	52208	54442	49974	2.26
	(9902)	(9355)	(10024)	[0.02]
Publications with Private Sector	12	17	7	5.07
Collaboration	(10)	(11)	(6)	[0.00]
State Private Sector Patents Per	162	167	157	0.69
1000 residents in 1981	(75)	(65)	(84)	[0.49]
State Per Capita Income in 1981	11118	11105	11132	-0.06
	(2203)	(2195)	(2236)	[0.95]
Observations	96	48	48	--

Notes: The sample contains one observation for each university in our sample. The main entries in columns (1) through (3) are the mean of the selected variable. The entries in parentheses in columns (1) through (3) are the standard deviation of the selected variables. Reported t-statistics are obtained from a regression of the selected variable on an indicator variable for universities in the above-median-university-research-expenditure category. All reported monetary amounts are in 1996 dollars. Also, all reported monetary amounts are in $1Million US. The only variables for which there are discrepancies in sample sizes are Patents and Patent Citations as not all universities patented in the initial year. Patents and Patent Citations have 58 observations total in 1981, 37 in High Expenditure and 21 in the Low Expenditure.

TABLE 2: The Effect of Stock Market Endowment Value Shock on Research Expenditure

Dependent Variable = Δ Research Expenditure

Model=	FD-OLS (1)	FD-OLS (2)
Δ Stock Index$_{t-1}$ x Initial Endowment	0.034***	0.027***
	(0.006)	(0.006)
Δ Stock Index$_{t-2}$ x Initial Endowment	--	0.018***
		(0.004)
Δ Stock Index$_{t-3}$ x Initial Endowment	--	0.004
		(0.006)
F-Statistic:		
Stock Index$_{t-1}$ x Initial Endowment	34.28	19.39
[P-Value]	[0.00]	[0.00]
Observations	1152	960

Notes: Source: Author's Calculations. The estimates presented are for two versions of equation (7) in the text. The unit of observation is the university-year level and the sample includes all 96 universities in the sample as described in the text and data appendix. The dependent variable is Δ Research Expenditure in year t. The main entries in columns (1) and (2) are coefficient estimates with each column representing a separate regression model with year fixed effects. The entries in parentheses in columns (1) and (2) are the standard errors of the coefficient estimates clustered at the university level.

* indicates significantly different from zero at the 10% level of significance
** indicates significantly different from zero at the 5% level of significance
*** indicates significantly different from zero at 1% level of significance

TABLE 3: The Effect of Research Expenditure on Basic Knowledge Production

Dependent Variable =	Publications	Δ Publications		Citations Per Publication	Δ Citations Per Publication	
Model =	OLS	FD-OLS	FD-IV	OLS	FD-OLS	FD-IV
	(1)	(2)	(3)	(4)	(5)	(6)
Research Expenditure$_{t-3}$	8.712***	--	--	0.008	--	--
	(0.656)			(0.005)		
Δ Research Expenditure$_{t-3}$	--	0.853*	8.993***	--	-0.004	-0.035***
		(0.447)	(2.049)		(0.003)	(0.009)
F-Statistic for First Stage	--	--	34.28	--	--	34.28
[P-value]			[0.00]			[0.00]
C-Stat (Difference-in-Sargan)	--	--	2.102	--	--	2.062
[P-value]			[0.15]			[0.15]
Observations	1152	1152	1152	1152	1152	1152
Baseline:						
Dependent Variable Mean		1017			10.7	
[Standard Deviation]		[684]			[4.5]	

Notes: Source: Source: Author's Calculations. The unit of observation is the university-year level and the sample includes all 96 universities in the sample as described in the text and data appendix. Columns (1) and (4) are for the model in equation (1). Columns (2) and (5) are for the model in equation (6). Columns (3) and (6) are for the model in equation (6), estimated by instrumental variables. The dependent variables Publications and Citations Per Publication are for year t. The main entries in columns (1) through (6) are coefficient estimates with each column representing separate regression model, all of which include year fixed effects. The entries in parentheses in columns (1) through (6) are the standard errors of the coefficient estimates clustered at the university level. The F-statistic is for the excluded instrument in equation (7). The C-Stat (Difference-in-Sargan) test statistic is for the difference between the FD-OLS and FD-IV estimates in columns (2) and (3), and (5) and (6) respectively.

* indicates significantly different from zero at the 10% level of significance
** indicates significantly different from zero at the 5% level of significance
*** indicates significantly different from zero at 1% level of significance

TABLE 4: The Effect of Research Expenditure on the Applied Knowledge Production

Dependent Variable =	Patents		Δ Patents		Citations Per Patent		Δ Citations Per Patent	
Model =	OLS	TOBIT	FD-OLS	FD-IV	OLS	TOBIT	FD-OLS	FD-IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Research Expenditure$_{t-3}$	0.075***	0.084***	--	--	0.005	0.009	--	--
	(0.016)	(0.019)			(0.003)	(0.005)		
Δ Research Expenditure$_{t-3}$	--	--	0.024	0.034	--	--	0.005	-0.112
			(0.016)	(0.100)			(0.013)	(0.071)
F-Statistic for First Stage	--	--	--	34.28	--	--	--	34.28
[P-value]				[0.00]				[0.00]
C-Stat (Difference-in-Sargan)	--	--	--	0.001	--	--	--	4.71
[P-value]				[0.92]				[0.03]
Observations	1152	1152	1152	1152	1152	1152	1152	1152
Baseline:								
Dependent Variable Mean			4.8				8.2	
[Standard Deviation]			(7.8)				(5.1)	

Notes: Source: Source: Author's Calculations. The unit of observation is the university-year level and the sample includes all 96 universities in the sample as described in the text and data appendix. Columns (1), (2), (5), and (6) are for the model in equation (1). Columns (3) and (7) are for the model in equation (6). Columns (4) and (8) are for the model in equation (6), estimated by instrumental variables. The dependent variables Patents and Citations Per Patent are for year t. The main entries in columns (1) through (8) are coefficient estimates with each column representing separate regression model, all of which include year fixed effects. The entries in parentheses in columns (1) through (8) are the standard errors of the coefficient estimates clustered at the university level. The F-statistic is for the excluded instrument in equation (7). The C-Stat (Difference-in-Sargan) test statistic is for the difference between the FD-OLS and FD-IV estimates in columns (2) and (3), and (5) and (6) respectively.

* indicates significantly different from zero at the 10% level of significance
** indicates significantly different from zero at the 5% level of significance
*** indicates significantly different from zero at 1% level of significance

TABLE 5: The Effect of Endowment Income on Other Research Funding Income

Dependent Variable =	Δ Federal Research Funding	Δ State Research Funding	Δ Private Industry Research Funding
Model =	FD-OLS	FD-OLS	FD-OLS
	(1)	(2)	(3)
	Panel A: Three Year Effect		
Δ Stock Index$_{t-4}$ x	-0.008*	-0.000	-0.001
Initial Endowment	(0.005)	(0.000)	(0.001)
Observations	1016	1016	1016
	Panel B: Five Year Effect		
Δ Stock Index$_{t-6}$ x	-0.011	0.001	-0.001
Initial Endowment	(0.008)	(0.001)	(0.001)
Observations	832	832	832
Baseline:			
Dependent Variable Mean	106	12	6
[Standard Deviation]	(103)	(18)	(7)

Notes: Source: Source: Author's Calculations. The unit of observation is the university-year level and the sample includes all 96 universities in the sample as described in the text and data appendix. Columns (1) and (3) are for the model in equation (7) with the Δ $R_{i,t-3}$ variable is replaced with the variable in the column header. The dependent variables Federal Research Funding, State Research Funding and Private Industry Research Funding are for year t. The main entries in columns (1) through (3) are coefficient estimates with each column representing separate regression model, all of which include year fixed effects. The entries in parentheses in columns (1) through (3) are the standard errors of the coefficient estimates clustered at the university level.

* indicates significantly different from zero at the 10% level of significance
** indicates significantly different from zero at the 5% level of significance
*** indicates significantly different from zero at 1% level of significance

TABLE 6: The Effect of Research Expenditure on Basic Science Production, By University Control Type

Dependent Variable =	Δ Publications		Δ Citations Per Publication	
Model =	FD-IV	FD-IV	FD-IV	FD-IV
Sample =	Public	Private	Public	Private
	(1)	(2)	(3)	(4)
Δ Research Expenditure$_{t-3}$	23.762*	12.680***	-4.694	-1.446***
	(13.869)	(3.467)	(3.393)	(0.564)
F-Statistic for First Stage	3.36	10.55	3.36	10.55
[P-value]	[0.07]	[0.00]	[0.07]	[0.00]
Observations	732	420	732	420
Baseline:				
Dependent Variable Mean	1318	1358	9.1	13.5
[Standard Deviation]	[821]	[1018]	[3.5]	[4.8]

Notes: Source: Source: Author's Calculations. The unit of observation is at the university-year level and the sample includes all 96 universities stratified by control type. Columns (1) through (4) are for the model in equation (6) estimated by instrumental variables. The dependent variables Publications and Citations Per Publication are for year t. The main entries in columns (1) through (6) are coefficient estimates with each column representing separate regression model, all of which include year fixed effects. The entries in parentheses in columns (1) through (6) are the standard errors of the coefficient estimates clustered at the university level. The F-statistic is for the excluded instrument in equation (7).

* indicates significantly different from zero at the 10% level of significance
** indicates significantly different from zero at the 5% level of significance
*** indicates significantly different from zero at 1% level of significance

TABLE 7: The Effect of Research Expenditure on Basic Science Production: Alternative IV Specifications

Model =	FD-IV	FD-IV	FD-IV	FD-IV
Specification=	Baseline: Endowment	Log Endowment	Top-Coded Endowment	Moving Average Endowment
	(1)	(2)	(3)	(4)
Panel A: IV Results *Dependent Variable = Δ Publications*				
Δ Research Expenditure$_{t-3}$	8.993*** (2.049)	7.043*** (1.681)	7.442*** (1.929)	7.084*** (1.681)
Baseline: Dependent Variable Mean [Standard Deviation]		1017 [684]		
Panel B: IV Results *Dependent Variable = Δ Citations Per Publication*				
Δ Research Expenditure$_{t-3}$	-0.035*** (0.009)	-0.014 (0.015)	-0.031** (0.013)	-0.047*** (0.009)
Baseline: Dependent Variable Mean [Standard Deviation]		10.7 [4.5]		
Panel C: First Stage Results *Dependent Variable = Δ Research Expenditure$_{t-3}$*				
Δ Stock Index$_{t-4}$ x Initial Endowment	0.034*** (0.006)	--	--	--
Δ Stock Index$_{t-4}$ x log (Initial Endowment)	--	4.789*** (1.286)	--	--
Δ Stock Index$_{t-4}$ x Max(95th Percentile, Initial Endowment)	--	--	0.046** (0.014)	--
Δ Stock Index$_{t-4}$ x Initial Endowment	--	--	--	0.027*** (0.006)
Δ Stock Index$_{t-5}$ x Initial Endowment	--	--	--	0.018*** (0.004)
Δ Stock Index$_{t-6}$ x Initial Endowment	--	--	--	0.004 (0.006)
F-Statistic for First Stage [P-value]	34.28 [0.00]	13.88 [0.00]	10.93 [0.00]	35.27 [0.00]
Observations	1152	1152	1152	960

Notes: Source: Source: Author's Calculations. The unit of observation is at the university-year level and the sample includes all 96 universities as described in the text and data appendix. Columns (1) through (4) contain four different specifications of the instrumental variables estimates. The table is broken into three panels, where Panel A and B present the instrumental variables estimates for the outcome indicated and Panel C presents the results of each first stage respectively. In Panels A and B, each cell represents a separate regression of the model in equation (6) where the first stage is given in Panel C. In Panel C each column represents a separate regression with coefficients presented for the given first stage specification. The entries in parentheses in columns (1) through (4) are the standard errors of the coefficient estimates clustered at the university level.

* indicates significantly different from zero at the 10% level of significance
** indicates significantly different from zero at the 5% level of significance
*** indicates significantly different from zero at 1% level of significance

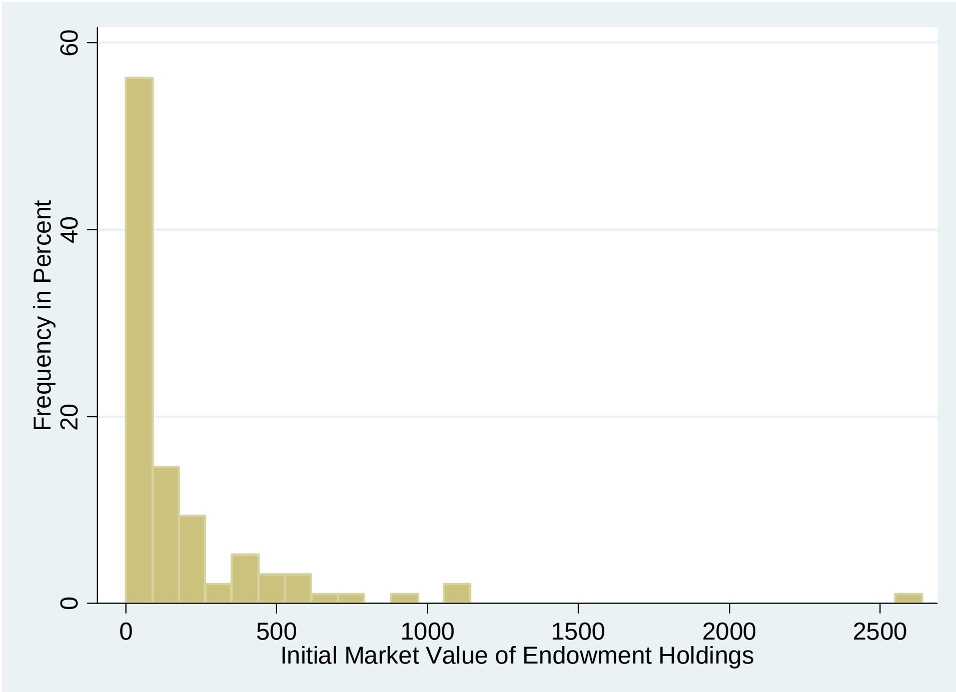
TABLE 8: The Effect of Research Expenditure on Basic Science Production: Identifying Assumption Validity Tests

Model =	FD-IV	FD-IV	FD-IV	FD-IV	FD-IV	FD-IV	FD-IV	FD-IV
					Initial Characteristic X Year Fixed Effects			
Specification =	Baseline	State Trends	University Trends	5-Year Lead	Papers Per Faculty	Number of Faculty	Tuition Per Student	Papers with Private Sector
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Dependent Variable = Δ Publications								
Δ Research $Expenditure_{t-3}$	8.993***	9.413***	12.211*	10.951**	11.192***	10.046***	10.414***	7.364***
	(2.049)	(2.205)	(6.059)	(4.323)	(3.812)	(1.869)	(1.892)	(2.547)
Δ Stock $Index_{t+2}$ x Initial Endowment	--	--	--	-0.041 (0.072)	--	--	--	--
Panel B: Dependent Variable = Δ Citations Per Publication								
Δ Research $Expenditure_{t-3}$	-0.035***	-0.017	-0.014	-0.008	-0.023	-0.031***	-0.040***	-0.036***
	(0.009)	(0.013)	(0.052)	(0.025)	(0.020)	(0.010)	(0.015)	(0.012)
Δ Stock $Index_{t+3}$ x Initial Endowment	--	--	--	-0.001 (0.001)	--	--	--	--
Added Fixed Effects	NONE	State	University	NONE	NONE	NONE	NONE	NONE
Added Interactions	NONE	NONE	NONE	NONE	YES	YES	YES	YES
F-stat for First Stage	34.28	26.86	4.25	7.77	23.51	32.64	24.35	23.51
[P-value]	[0.00]	[0.00]	[0.04]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Observations	1152	1152	1152	960	1152	1152	1104	1080
Baseline:								
Dependent Variable Mean				1017				
[Standard Deviation]				[684]				

Notes: Source: Source: Author's Calculations. The unit of observation is at the university-year level and the sample includes all 96 universities as described in the text and data appendix. The table is broken into two panels where Panel A and B present the instrumental variables estimates for the outcome indicated. Columns (1) through (8) report instrumental variables estimates where each cell represents a separate regression of the model in equation (6). The specification in each column adds a specific control or interaction to the baseline model in (6) as indicated by *Specification* and noted in the text. The main entries in columns (1) through (8) are coefficient estimates and the entries in parentheses in columns (1) through (8) are the standard errors of the coefficient estimates, clustered at the university level. The F-statistic is for the excluded instrument in equation (7).

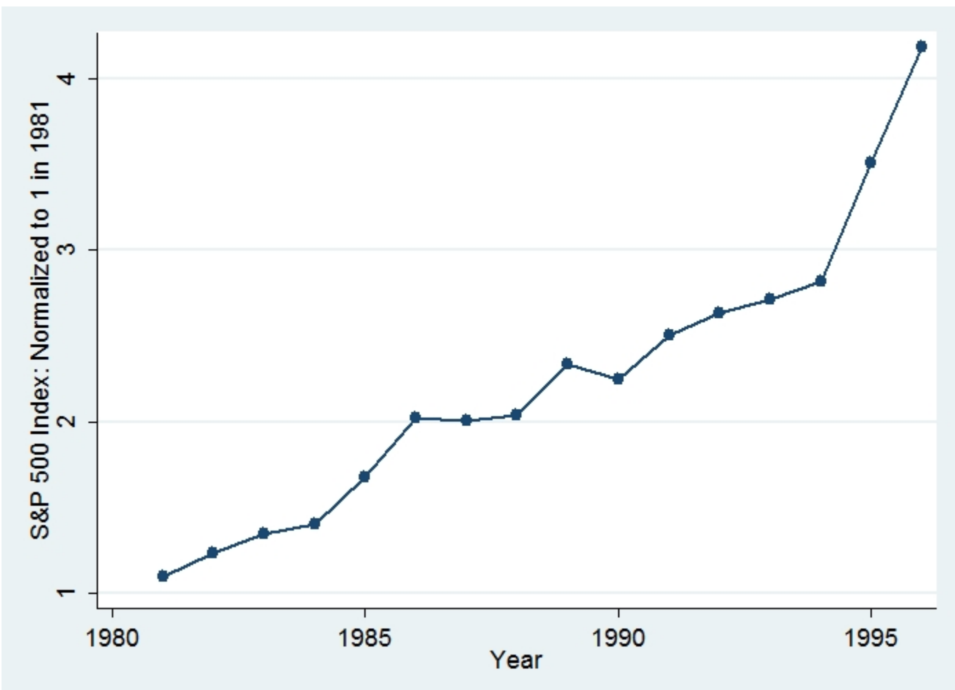
* indicates significantly different from zero at the 10% level of significance
** indicates significantly different from zero at the 5% level of significance
*** indicates significantly different from zero at 1% level of significance

FIGURE 1: Cross-Sectional Distribution of Initial Endowment Market Values



Source: Authors' Calculations from HEGIS 1981 Data

FIGURE 2: Standard and Poor's 500 Stock Index: 1981-1996



Source: Authors' Calculations from Standard and Poor's 500 stock index data. The stock index is normalized to be one in 1981.

1 Data Appendix (Not For Publication)

Our study analyses the knowledge production-funding relationship for research producing Universities throughout the United States. This data appendix outlines each data sources we utilize in the analysis, and how our sample of 96 universities is constructed.

1.1 Data Sources

Our data sources can be broken into four primary categories: outcomes, university financial data, university characteristics, and finally other data such as price deflators or stock market indices. The first three categories are all measured at the university level, while other variables are measured at various levels of aggregation as noted.

1.1.1 Outcomes

NBER-Rensselaer Polytechnic Institute Scientific Papers Database (NBER-RES) The NBER-RES database (Adams and Clemons, 2009) contains information on the academic output of the top 110 non-profit institutions and 200 U.S. RD performing firms. The data covers all publications and the citations to those publications for the 310 institutions from 1981 through 1996. The data measures all publications, citations and collaboration at the institution-field-year level. We aggregate the data to university-year level totals, as our research expenditure measures are at the university-year level. The fact that the data is first available in 1981 determines the initial year in our analysis. The data is available at http://www.nber.org/RPI-sci-pap/.

NBER-Patent Citations Data File (NBER-PAT) The NBER-PAT database (Hall, Jaffe, and Trajtenberg, 2001) contains data on 3,190,796 patents granted between January 1963 and December 2002. In addition to the date the patent is applied for, the date received, and the identity of the assignee, as well as future citation and collaboration information. The data is available at the individual patent level, but can also be pooled by assignee, industry, or region. For our purposes, we keep only patents whose assignee is a university in the US. We identify university assignees with a crosswalk developed by developed by the United States Patent and Trademark Office which maps universities to the assignees in the NBER-PAT. In addition, we drop any patent for which the primary assignee (first assignee listed) is not a university. We then matched this NBER-PAT sample to our university sample, and are left with 34,675 patents out of the total, or just over 1% of the total number of patents produced. Citation and collaboration data are directly tied to these 34,675 patents. The patent year is measured at the application date rather than grant date as we are interested in the date closest to the actual innovation. The data is available at http://www.nber.org/patents/.

We face one data constraint in attributing patents that are only recorded at the university system level to individual campuses within the system for two large systems: the University of California System and the University of New York System. For these two systems we assign the systemwide patent data to the individual campuses in the following way. We simply use a within system academic publication share from the NBER-RES to apportion the patent data to the individual campuses. It is also worth noting that not all of the universities in our sample have recorded patents. Six of the 96 universities record zero patents over the time period.

1.1.2 University Financial and Characteristic Data

All university level data were retrieved from two sources. First, the Higher Education General Information Survey (HEGIS) is used to compile financial and descriptive data on universities from 1981 through 1983. Then from 1984 through 1996 the Integrated Post-Secondary Education Data System (IPEDS) is used to complete the data series, all at the university-year level. The HEGIS data is available at http://www.icpsr.umich.edu/icpsrweb/ICPSR/ and the IPEDS data is available at http://nces.ed.gov/ipeds/.

University Micro Data Detailed Both the HEGIS and IPEDS include rich data on everything from student counts and tuition through detailed line-item budgetary expenditures. We extract all of our measures of expenditure, income, student counts, faculty counts, and faculty salaries from the data. The university fiscal year runs from July of the previous year through June of the current year. All expenditure and revenue items are reported on a fiscal year basis.

We encounter one missing data issue with respect to university research spending. Unfortunately, 10 of the universities do not report research spending in the final year of our sample period. To keep these universities in the sample we used the national growth rate in research spending between 1995 and 1996 to impute research spending for these 10 universities in 1996 from their 1995 data. The 10 universities are: California Institute of Technology, Case Western Reserve University, Cornell University, Duke University, Emory University, John's Hopkins University, Princeton University, University of California-Irvine, University of Iowa, University of Southern California.

1.1.3 Other Data Sources

Census Data State level income and college graduate percentages from 1980 Decennial Census are obtained from the Historical, Demographic, Economic, and Social Data: The United States, 1790-2000 file. This data is available at http://www.icpsr.umich.edu/ icpsrweb/ICPSR/studies/2896?archive=ICPSRq=haines.

NSF Survey of Federal SE Support We use the NSF Data on industry based grant funding to measure industry based grants to universities. The data is available at http://www.nsf.gov/statistics/showsrvy.cfm?srvy_ CatID=4& srvy_ Seri=12

Stock Market Index Yearly averages were generated from the monthly Standard and Poors 500. The historical prices were retrieved from http://finance.yahoo.com/q/ hp?s=ĜSPC. As universities fiscal year budgetary planning occurs at a one year lag from stock market fluctuations, the index year is lagged one year. For example, in the analysis, the 1981 SP 500 data is actually the average over 1980 to account for the budgetary lag. Also, the years are matched on the July to June basis to match the fiscal year in the HEGIS and IPEDS data. Finally, the index is normalized to one in the base year; 1981.

Price Deflator All financial data is deflated to be 1996 constant dollars utilizing the Quantity and Price Indexes for Gross Domestic Product available from the Bureau of Economic Statistics at http://www.bea.gov/national/nipaweb/
TableView.asp?SelectedTable=13&Freq=Qtr& FirstYear=2007& LastYear=2009.

NBER-Patent Citations Data File to University Crosswalk The crosswalk is utilized to map the patent data to HEGIS and IPEDS. The data is available at
http://www.uspto.gov/web/offices/ac/ido/oeip/taf/univ/univ_ toc.htm.

1.2 Sample Construction

The initial sample consists of the top 110 research universities as defined in the NBER-Rensselaer Polytechnic Institute Scientific Papers Database (NBER-RES) (Adams and Clemons, 2008). The NBER-RES database contains information on publication outcomes for all researchers at leading researchers in the United States. The universities in the sample account for 85 % of all the knowledge produced by universities in the United States (Adams and Clemons, 2008).

Our central analysis is concerned with large research universities that offer undergraduate and graduate degrees in multiple fields. In addition, those institutions that do not offer any degrees do not typically in the universe for the HEGIS/IPEDS data. We thus first drop any institutions in the NBER-RES database that are focused solely on medical or other narrowly focused areas, such as oceanographic institutes. Specifically we drop the following institutions: University of Texas Houston Health Science Center, Woods Hole Oceanographic Institute, University of Texas San Antonio Health Science Center, University of Texas Southwestern Medical Center of Dallas, Oregon Health Sciences University, Baylor College of Medicine. This leaves us with a total of 104 research universities in the initial sample.

Our sample of 104 universities is then matched to the institutions in the Higher Education General Information Survey (HEGIS) and Integrated Post-Secondary Education Data System (IPEDS) data. Two data constraints result in a further 8 schools being dropped from the sample. First, as research spending for Rutgers University is missing for a large number of years, and because no sensible imputation can be utilized, it is dropped. There is also a small set of universities are missing the endowment market values in the base year, which is necessary for the instrument generation, and are therefore dropped. These institutions are: University of Connecticut, University of Kansas Main Campus, State University of New York at Buffalo, Baylor University, University of Utah, Virginia Commonwealth University. The second data constraint we face is that the observations for University of Texas - Austin appear to report endowment market values for the entire system, which includes nine universities and six health centers. As the other University of Texas campuses and health centers are outside of the sample scope, we drop this observation as well. This completes the sample construction, leaving us with 96 universities in the final sample for analysis.

TABLE A1: Variable Definitions and Sources

Variable	Definition and Creation Notes	Source
(1) Output:		
Papers	Publication counts at the university level	NBER-PAP
Citations Per Paper	Future Citations on all papers published by researchers at a university in reference year	NBER-PAP
Patents	Patent counts at the university level	NBER-PAT
Citations Per Patent	Future Citations on all patents granted to researchers at a university in reference year	NBER-PAT
(2) University Expenditure, Grants and Characteristics:		
Research Expenditure	Sum of directed research (both restricted and unrestricted) and instruction funds. This is then weighted by the fraction of time spent on research activities as reported by the National Study of Postsecondary Faculty available at: http://nces.ed.gov/surveys/nsopf/	HEGIS/IPEDS
Total Expenditure	Total Current Funds Expenditure	HEGIS/IPEDS
Endowment Market Value in 1981	The Market Value of the Endowment in 1981	HEGIS/IPEDS
Federal Research Grants	Federal Grant Funds - Unrestricted	HEGIS/IPEDS
State &Local Research Grants	State and Local Grant Funds - Unrestricted	HEGIS/IPEDS
Private Industry Research Grants	Industry Source Grant Funds	NSF
Number of Students	Total Graduate and Undergraduate Students	HEGIS/IPEDS

Number of Faculty	Total faculty	HEGIS/IPEDS
Public	Binary indicator of University Publically Controlled	HEGIS/IPEDS
U.S. News Quality Ranking	An ordinal ranking of college quality on a scale of 1 to 4, 1 being lowest 4 being highest based the U.S. News and World Report college rankings in 1991	U.S. News 1991
Mean Faculty Salary	Average Faculty Salary	HEGIS/IPEDS
Papers with Private Sector	Count of papers in which the primary assignee is a university and the collaborating entity is a private firm.	NBER-PAT
State Per Capita Income in 1981	Per Capita Income in University Location State in 1981	City County Data Book
State Private Sector Patents Per 1000 residents in 1981	Patent Counts in 1981 per 1000 Population in 1980	NBER-PAT / City County Data Book

(3) Other Variables:

Stock Index	S&P 500 Index at the annual level where 1981 is the base year and the index is normalized to equal one. In addition, the 1981 index is the average from June of 1980 through July of 1981 in order to matches the yearly expenditures as reported in the HEGIS/IPEDS by fiscal year.	finance.yahoo.com

NBER-PAP is the Scientific Papers Database obtained through the publicly accessible data sources through the National Bureau of Economic Research (NBER): http://www.nber.org/RPI-sci-pap/. The NBER-PAT is the Patent Citations Data File obtained through the publicly accessible data sources through the NBER: http://www.nber.org/patents/. The HEGIS/IPEDS is a database on all university financial, faculty, and enrollment variables located through the Inter-University Consortium of Political and Social Research (ICPSR): http://www.icpsr.umich.edu/icpsrweb/ICPSR/.

TABLE A2: List of Universities in Sample

Arizona State University	Princeton University	University of Kentucky
Boston University	Purdue University Main Campus	University of Maryland Baltimore County
Brandeis University	Rice University	University of Maryland College Park
Brown University	Rockefeller University	University of Massachusetts
California Institute of Technology	Stanford University	University of Miami
Carnegie Mellon University	State University of New York	University of Michigan
Case Western Reserve University	Syracuse University Main Campus	University of Minnesota
City University of New York	Texas A & M University	University of Missouri
Colorado State University	Tufts University	University of Nebraska
Columbia University	Tulane University	University of New Hampshire
Cornell University	University of Alabama, The	University of New Mexico Main Campus
Dartmouth College	University of Alaska Fairbanks	University of North Carolina at Chapel
Duke University	University of Arizona	University of Oregon
Emory University	University of California-Berkeley	University of Pennsylvania
Florida State University	University of California-Davis	University of Pittsburgh, Pittsburgh Ca
Georgia Institute of Technology	University of California-Irvine	University of Rochester
Harvard University	University of California-Los Angeles	University of Southern California
Indiana University	University of California-Riverside	University of Tennessee
Iowa State University	University of California-San Diego	University of Vermont
Johns Hopkins University	University of California-San Francisco	University of Virginia
Lehigh University	University of California-Santa Barbara	University of Washington
Louisiana State University	University of California-Santa Cruz	University of Wisconsin
Loyola University of Chicago	University of Chicago	Utah State University
Massachusetts Institute of Technology	University of Cincinnati Main Campus	Vanderbilt University
Michigan State University	University of Colorado at Boulder	Virginia Polytechnic Institute
New Mexico State University	University of Delaware	Wake Forest University
New York University	University of Florida	Washington State University
North Carolina State University	University of Georgia	Washington University
Northwestern University	University of Hawaii at Manoa	Wayne State University
Ohio State University	University of Illinois at Chicago	West Virginia University

Oregon State University	University of Illinois at Urbana-Champaign	Yale University
Pennsylvania State University	University of Iowa	Yeshiva University